AD-758 224

NETWORK MULTIPLE STATIONS DISCRIMINANT FUNCTIONS

R. Shumway, et al

Teledyne Geotech

Prepared for:

Air Force Technical Applications Center
Advanced Research Projects Agency

3 November 1972

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

NETWORK MULTIPLE STATION DISCRIMINANT FUNCTIONS

R. H. SHUMWAY
CONSULTANT TO THE SEISMIC DATA LABORATORY
R. R. BLANDFORD
SEISMIC DATA LABORATORY

3 NOVEMBER 1972

Prepared for
AIR FORCE TECHNICAL APPLICATIONS CENTER
Washington, D.C.

Under
Project VELA UNIFORM

Sponsored by
ADVANCED RESEARCH PROJECTS AGENCY
Nuclear Monitoring Research Office
ARPA Order No. 1714

TELEDYNE GEOTECH
ALEXANDRIA LABORATORIES

Neither the Advanced Research Projects Agency nor the Air Force Technical Applications Center will be responsible for information contained herein which has been supplied by other organizations or contractors, and this document is subject to later revision as may be necessary. The views and conclusions presented are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency, the Air Force Technical Applications Center, or the U S Government.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) Teledyne Geotech Alexandria, Virginia	2a REPORT SECURITY CLASSIFICATION Unclassified 2b GROUP

3 REPORT TITLE

MULTIPLE STATION DISCRIMINANT FUNCTIONS

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific

5 AUTHOR(S) (Last name, first name, initial)

Shumway, R.; Blandford, R.R.

6. REPORT DATE 3 November 1972	7a TOTAL NO. OF PAGES 35	7b NO OF REFS 7

8a CONTRACT OR GRANT NO. F33657-72-C-0009 b. PROJECT NO VELA T/2706 c. ARPA Order No. 1714 d. ARPA Program Code No. 2F-10	9a ORIGINATOR'S REPORT NUMBER(S) 299 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10 AVAILABILITY/LIMITATION NOTICES

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

11 SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency Nuclear Monitoring Research Office Washington, D.C.

13 ABSTRACT

An extension is made of linear discriminant analysis to the case where multiple station observations are available for each event. Multivariate regression is used to estimate the mean vectors and covariance matrix in the multiple station discriminant function. The same stations need not be observed for all events as a separate discriminant vector is derived for each station observation and only the discriminant functions available for each event are added. The identification curves for this multiple station discriminant are calculated for surface body wave pairs from a population of 20 earthquakes and 9 presumed explosions observed at various sub-combinations of six LRSM stations. The results obtained indicate that the multiple station discriminant function is superior to the usual method which treats the mean vector for each event as a single observation, a result of importance in the application of discriminants by a network of stations.

14 KEY WORDS

Seismic Discrimination
Multivariate Regression
Multiple Station Discriminants

Unclassified

Security Classification

MULTIPLE STATION DISCRIMINANT FUNCTIONS

SEISMIC DATA LABORATORY REPORT NO. 299

AFTAC Project No.:	VELA T/2706
Project Title:	Seismic Data Laboratory
ARPA Order No.:	1714
ARPA Program Code No.:	2F-10
Name of Contractor:	TELEDYNE GEOTECH
Contract No.:	F33657-72-C-0009
Date of Contract:	01 July 1971
Amount of Contract:	$ 2,482,460
Contract Expiration Date:	30 June 1973
Project Manager:	Robert R. Blandford (703) 836-3882

P. O. Box 334, Alexandria, Virginia 22314

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In an analysis of seismic data made for the purpose of discriminating between earthquakes and explosions, a common practice (Booker and Mitronovas (1964), Capon, et. al. (1967), Ericsson (1970), and Shumway and Blandford (1970)) is to use the mean of a seismic discriminant measured at several stations as the basic observation. For example, surface wave and body wave pairs plotted in two dimensions may actually be mean values deduced from differing numbers or configurations of stations. The discriminant analysis which results ignores the fact that the measurements may have different means due to differing station configurations and will surely have different variances because of the different sample sizes over which each mean has been determined. For these reasons, one would prefer to work with the values observed at each of the stations which record the event.

This report will describe and evaluate a method for handling the discriminant analysis when observations are made at more than one station. A linear criterion which is an extension of the standard discriminant function will be considered. Evaluation will be in terms of standard operating characteristic curves (termed "identification curves" by Ericsson, (1970) which measure the predicted explosion detection capability as a function of the false alarm rate. Ericsson, (1970) contains a clearly written explanation of the political ramifications leading to the use of identification curves as well as an excellent review of the seismic discrimination literature.

2. LINEAR DISCRIMINANT ANALYSIS WITH MULTIPLE OBSERVATIONS

In the standard situation a p x 1 vector of proposed discriminants, $\underline{X}' = (X_1,\ldots,X_p)$, is observed and we wish to determine whether the vector came from a population of earthquakes (hypothesis H_1) or a population of explosions (hypothesis H_2). In general, multivariable normality of the vector $\underline{X}$ is assumed with only the mean value vectors for the two populations allowed to vary. The effect of variations likely to be observed in the covariance matrices seems to be small for seismic data (Shumway and Blandford, 1970). In the case where N_s stations record observations from the same event, we would like to assign each station a different mean value. Then, one has available a sample of N_s vector discriminants ($\underline{X}_j$, j = 1, 2,...,N_s) for classifying the events as belonging to either the earthquake or explosion group. Under H_1 (earthquakes),($\underline{X}_j$, j = 1, 2,...N_s) are N_s independent normal vectors with mean ($\underline{V}_{1j}$, j = 1,...,N_s) and covariance matrix Σ while under H_2 they are independent normal vectors with means ($\underline{V}_{2j}$, j = 1,...,N_s) and identical covariance matrix Σ. An application of the Neyman-Pearson argument to this case yields the discriminant function:

$$d(\underline{X}_1,\ldots,\underline{X}_{N_s}) = \sum_{j=1}^{N_s} \underline{d}_j' \underline{X}_j - \sum_{j=1}^{N_s} (CT)_j \quad (1)$$

where

$$\underline{d}_j' = (\underline{V}_{1j} - \underline{V}_{2j})' \Sigma^{-1} \quad (2)$$

and

$$(CT)_j = 1/2(\underline{V}_{1j} + \underline{V}_{2j})' \textstyle\sum^{-1} (\underline{V}_{1j} - \underline{V}_{2j}) \tag{3}$$

Equations (1) through (3) demonstrate that the discriminant function for observations from more than one station is just the sum of the ordinary discriminant functions evaluated at the station means. It may be that general linear procedures of the form

$$a(\underline{X}_1,\ldots,\underline{X}_{N_s}) = \sum_{j=1}^{N_s} \underline{a}_j' \underline{X}_j \tag{4}$$

might be useful discriminants. An example of a linear function which is not necessarily the linear discriminant function would be $M_{sj} - 1.9\, m_{bj}$ where the vector $\underline{X}_j' = (M_{sj}, m_{bj})$ is a surface wave body wave pair and $\underline{a}_j' = (1, -1.9)$. We note parenthetically that for $N_s = 1$, or for equal station means, equations (1) through (3) reduce to the usual equations for the linear discriminant function, so that the procedure is equivalent to discrimination using network averages.

In order to apply equations (1) through (3) we must either know $\underline{V}_{1j}$, $\underline{V}_{2j}$, $\sum$, $j = 1,\ldots,N_s$, or obtain consistent estimators for these terms.

In Section 4 we discuss the usual multivariate theory as it applies to estimating the generalized means and covariance matrix.

3. THE IDENTIFICATION CURVE AS A MEASURE OF PERFORMANCE

For discriminant functions of the form (1) an expected future performance can be calculated using the result that $d(\underline{X}_1,...,\underline{X}_{N_s})$ is a normal random variable with mean $\sum_j D_j^2$ under H_1 (earthquake), and $-\sum_j D_j^2$ under H_2 (explosion), and a variance $2\sum_j D_j^2$ under both hypotheses, with

$$D_j^2 = 1/2(\underline{V}_{1j} - \underline{V}_{2j})'\Sigma^{-1}(\underline{V}_{1j} - \underline{V}_{2j}) \tag{5}$$

In general we wish to control the probability of a false alarm. A false alarm occurs when an earthquake is incorrectly called an explosion. That is when H_1 is true and the value of the discriminant function $d(\underline{X}_1,...,\underline{X}_{N_s})$ is low enough to cause the observation to be classified with the explosion population. If α is the false alarm probability and $C(\alpha)$ is a constant depending only on α we accept H_2 (explosion) if

$$d(\underline{X}_1, ..., \underline{X}_{N_s}) \leq C(\alpha).$$

For a given false alarm probability

$$\alpha = Pr_{H_1} (d(\underline{X}_1,\ldots,\underline{X}_{N_s}) \leq C(\alpha))$$

$$= \Phi \left(\frac{C(\alpha) - \sum D_j^2}{2\sum D_j^2}\right) \tag{6}$$

with Φ denoting the cumulative normal distribution function, equation (6) may be solved for $C(\alpha)$ yielding

$$C(\alpha) = (\sum_j D_j^2)\ (1 + 2\ \Phi^{-1}(\alpha)) \tag{7}$$

where Φ^{-1} denotes the inverse of the cumulative normal distribution function. The probability of correctly identifying the explosion,

$$P_d(\alpha) = Pr_{H_2} (d(\underline{X}_1,\ldots,\underline{X}_{Ns}) \leq C(\alpha))$$

$$= \Phi \left(\frac{C(\alpha) + \sum D_j^2}{2\sum D_j^2}\right) \tag{8}$$

can be calculated as a function of the false alarm rate α.

We note that for a general linear function of the form (4) with means μ_1, $\mu_2 < \mu_1$, and variances σ_1^2, σ_2^2, equations (7) and (8) become

$$C(\alpha) = \mu_1 + \sigma_1 \Phi^{-1}(\alpha) \tag{9}$$

and

$$P_d(\alpha) = \Phi \left(\frac{C(\alpha) - \mu_2}{\sigma_2}\right) \tag{10}$$

If $\mu_2 > \mu_1$ the equations become

$$C'(\alpha) = \mu_1 + \sigma_1 \Phi^{-1} (1-\alpha) \tag{11}$$

and

$$P_d'(\alpha) = 1 - \Phi \left(\frac{C(\alpha) - \mu_2}{\sigma_2}\right) \tag{12}$$

A plot of $P_d(\alpha)$, $P_d'(\alpha)$ against α gives a measure of explosion detection capability and has been termed an identification curve by Ericsson (1970).

A hypothetical identification curve is shown in Figure 1. In general, Ericsson postulates that an acceptable operating region is determined by keeping the false

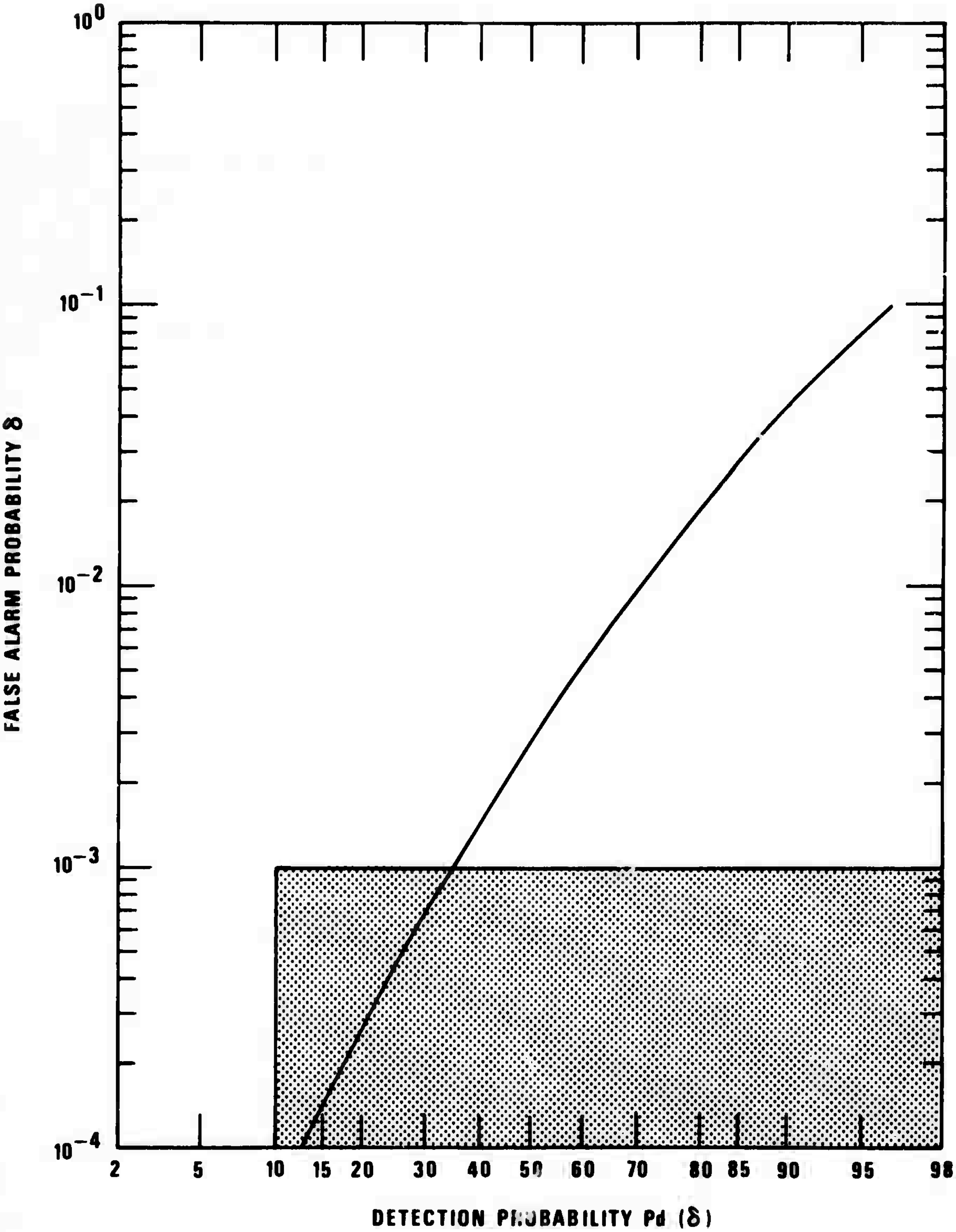

Figure 1. A hypothetical identification curve showing the region which combines an acceptable detection probability $Pd>.10$ with an acceptable risk $\alpha < 10^{-3}$.

6 b

alarm probability at 1×10^{-3} or less while maintaining the desired detection probability at .10 or greater. Figure 1 gives a geometric interpretation to the politically acceptable region along with a hypothetical identification curve showing that a maximum of 35% of the explosions could be detected within the acceptable false alarm probability of 1×10^{-3}.

Figure 1 suggests that two proposed procedures leading to normally distributed discriminants can be compared by examining their respective identification curves. A procedure is superior to another if its detection curve is located to the right of the other procedure over the region of interest.

4. ESTIMATION OF THE STATION MEANS AND COVARIANCE MATRIX FOR EARTHQUAKES AND EXPLOSIONS

In order to estimate the unknown parameters $\underline{V}_{1j}$, $\underline{V}_{2j}$, Σ, $j = 1,\ldots,N_s$ we suppose that multiple station observations are available for N_1 earthquakes and N_2 explosions on possible discriminants. Denote the k'th $p \times 1$ vector of discriminants measured at station j for an event of type i by $\underline{y}_{ijk}$ so that under H_1 (earthquake) $\underline{y}_{1jk}$ has a normal distribution with mean vector $\underline{V}_{1j}$ and covariance matrix Σ while under H_2 (explosion) $\underline{y}_{2jk}$ has a mean vector $\underline{V}_{2j}$ and covariance matrix Σ, $j = 1, 2,\ldots,J$; $k = 1, 2,\ldots, N_{ij}$. N_{1j} is the number of observations measured by the j'th station for an earthquake and N_{2j} is the number of observations made by the j'th station for an explosion.

In this case it is well known that the maximum likelihood estimators for the mean vectors are

$$\hat{V}_{ij} = N_{ij}^{-1} \sum_{k=1}^{N_{ij}} \underline{y}_{ijk}, \quad i = 1, 2,\ j = 1,\ldots,J \tag{13}$$

and that an unbiased estimate for the covariance matrix which is proportional to the maximum likelihood estimator is

$$\hat{\Sigma} = \Big(\sum_{ij}(N_{ij} - 1)\Big)^{-1} \sum_{ijk}(\underline{y}_{ijk} - \hat{\underline{V}}_{ij})(\underline{y}_{ijk} - \hat{\underline{V}}_{ij})' \tag{14}$$

In practice the computations are rather involved for the above expression and one may wish to try different models and test various hypotheses about the vectors $\underline{V}_{ij}$. Therefore, it is convenient to formulate all models in terms of the multivariate linear hypothesis structure described below.

Let $\underline{y}_k$, $k = 1,\ldots,N$ be $p \times 1$ independently distributed normal vectors with means $B'\underline{x}_k$ and covariance matrix Σ. Here B is a $q \times p$ matrix of unknown parameters, and $\underline{x}_k$, $k = 1,\ldots,N$ are fixed known $q \times 1$ vectors. Then the Nxp matrix $Y = (\underline{y}_1,\ldots,\underline{y}_N)'$ has mean value XB, where $X = (\underline{x}_1,\ldots,\underline{x}_N)'$. For this ca e the maximum likelihood estimate for B is

$$\hat{B} = (X'X)^{-1} X'Y \tag{15}$$

while an unbiased estimate proportional to the maximum likelihood estimator for Σ is

$$\hat{\Sigma} = (N-q)^{-1} (Y-X\hat{B})'(Y-X\hat{B}) \tag{16}$$

For this setup, the program from Dixon (1970) yields $\hat{B}$ and $\hat{\Sigma}$, as well as the results of testing hypotheses of the form

$$ABC' = D \tag{17}$$

where $A(r \times q)$, $B(q \times s)$, and $D(r \times s)$ are matrices specified to generate a designated hypothesis. The matrix C must be $(s \times s)$. We illustrate the procedure in the next section.

5. EXAMPLE

The procedure implied by the theoretical approach in Sections 2, 3, and 4 will be illustrated using 20 earthquakes and 9 explosions observed at various sub-combinations of six recording stations. The observations shown in Table I are on body wave (m_b) and surface wave (M_s) magnitudes. We note that most events are not recorded at all stations. Hence, it would be useful to estimate an expected explosion detection probability for each sub-combination of stations which can be expected to appear as an observation.

We assume as in the previous section that the k'th 2 x 1 vector $\underline{y}_{ijk} = (M_{sijk}, M_{bijk})'$ pair at station j = 1, 2,...,6 for event i = 1, 2 (earthquake or explosion) has the 2 x 1 mean vector:

$$\underline{V}_{ij} = \begin{pmatrix} v_{ij1} \\ v_{ij2} \end{pmatrix} \tag{18}$$

where components denote respectively the theoretical mean of the M_s and m_b measurements at station j for an event of type i. In general, the estimates for $\underline{V}_{ij}$ and Σ are most important as they are needed for the multiple station discriminant function given in Section 2.

TABLE I

M_s-m_b Pairs for 29 Events
Recorded at 6 LRSM Stations

Stations	NP		WH		PG		RK		HN		KN	
Earthquakes (1)	m_b	M_s	m_b	M_s	m_b	M_s	m_b	M_s	m_b	M_s	MB	M_s
EQ1	6.25	4.30					5.07	4.20	5.99	4.94		
EQ2	5.77	4.23					4.80	3.81	4.86	3.75		
EQ3	6.26	6.33							6.33	6.39	6.35	6.19
EQ4	6.24	4.15					5.24	4.70	4.61	4.62		
EQ5	6.46	6.17			6.40	6.29	6.04	6.57	5.62	6.46	5.29	6.46
EQ6	6.21	4.59			4.86	4.52	5.01	4.36	4.98	4.20		
EQ7	5.45	5.01			5.07	5.38					4.98	4.91
EQ8	5.56	4.41	4.52	4.74	4.61	4.28	4.47	4.56			4.61	4.51
EQ9	5.89	5.36	5.22	5.67	5.30	5.42	5.43	5.48	5.36	5.62	4.96	5.39
EQ10	5.44	4.93					4.66	4.30	4.94	4.21		
EQ11			5.22	4.40	5.30	4.22	5.23	4.62			5.31	4.40
EQ12	5.85	4.69	5.18	4.73	5.40	4.50	5.31	4.73	5.24	4.77	4.62	4.80
EQ13	5.37	3.58			4.62	3.67	5.13	3.58	4.78	3.81		
EQ14	5.81	3.86	4.92	3.90	4.70	4.08	5.18	4.10	4.95	4.01		
EQ15	6.05	4.51	4.98	4.93	4.85	4.84	4.89	4.75	4.80	4.52	4.78	4.75
EQ16	5.90	4.86	5.06	5.05	4.94	4.90	4.83	4.81	4.90	4.84	4.85	4.95
EQ17	5.81	4.94			4.95	4.75	5.15	4.95	4.72	4.57		
EQ18			4.51	4.36	4.54	3.95	4.63	3.92				
EQ19	5.98	5.00	4.83	5.17	4.72	4.91	4.67	5.02	5.21	4.89	4.60	4.65
EQ20			5.51	5.40	5.48	5.35	5.23	5.45	5.78	5.33	5.17	5.11
Presumed Explosions (2)												
EX1							6.19	4.10	5.81	4.58	6.12	4.31
EX2	6.13	4.33			6.08	4.45	5.99	4.40	5.75	4.25	5.93	4.99
EX3	6.25	4.32			6.08	4.45	5.99	4.40			5.73	4.26
EX4	6.02	3.84	6.25	3.76	5.84	4.02	5.75	4.38	5.37	4.33	5.72	4.11
EX5							5.91	3.82	5.82	4.00		
EX6	6.33	4.31			6.09	4.75	6.90	4.79	6.71	4.65	6.17	4.46
EX7					5.68	4.19	6.69	4.46	6.01	4.24		
EX8	6.14	4.28			5.68	4.12	6.59	4.46	6.33	3.93	5.76	3.98
EX9	6.28	4.14					6.85	4.62	6.39	4.19		

Of more than incidental interest are the hypotheses listed below.

In general are earthquake and explosion mean vectors equal?

$$H: \sum_{j=1}^{6} \underline{V}_{1j} = \sum_{j=1}^{6} \underline{V}_{2j} \tag{19}$$

Are earthquake and explosion means equal on a station-by-station basis?

$$H: \underline{V}_{1j} = \underline{V}_{2j},\ j = 1,\ldots,6 \tag{20}$$

This also gives simultaneous confidence intervals for contrasts comparing earthquakes and explosions on a station-by-station basis.

How significant is the linear contrast $M_s - 1.9\ m_b$ on a station-by-station basis?

$$H: -1.9V_{1j1} + V_{1j2} = -1.9V_{2j1} + V_{2j2},\ j = 1,\ 2,\ldots,6 \tag{21}$$

For the input to program BMD-X-63,Dixon (1970), we note that the 12 x 2 parameter matrix B is given by

$$B = \begin{bmatrix} V_{111} & V_{112} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ V_{161} & V_{162} \\ V_{211} & V_{212} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ V_{261} & V_{262} \end{bmatrix} = \begin{bmatrix} \underline{V}'_{11} \\ \cdot \\ \cdot \\ \cdot \\ \underline{V}'_{16} \\ \underline{V}'_{21} \\ \cdot \\ \cdot \\ \cdot \\ V'_{26} \end{bmatrix} \qquad (22)$$

The 123 x 12 matrix X has in each row a vector which generates an observation in Table I. For example, the component which generates the observation at PG for the 16th quake, say (4.94, 4.90), should have mean $\underline{V}'_{13}$ so that the row vector for X should be (0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0) and similarly for the other observations in the table. In order to generate hypotheses (19) through (21), it is sufficient to note that they are of the form ABC' = D where (19), for example, has

$$A = (1, 1, 1, 1, 1, 1, -1, -1, -1, -1, -1, -1,),$$

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad D = (0, 0).$$

The generalized linear hypotheses program yielded the estimate for B given in Table II. The table is in the format of equation (22). The table tends to indicate that there are some significant differences between the means. The standard errors are computed by noting that the program produces an estimate for the covariance matrix

$$\hat{\Sigma} = \begin{pmatrix} .16 & .15 \\ .15 & .38 \end{pmatrix} \tag{23}$$

and that the variance covariance matrix of the rows of $\hat{B}$ is given by $\Sigma \otimes (X'X)^{-1}$ where $\otimes$ denotes the Kronecker product, and

$$(X'X)^{-1} = \text{Diag}\,(.06, .10, .07, .06, .06, .09, .17, 1.00, .14, .11, .12, .17). \tag{24}$$

TABLE II

Estimated Mean M_s, m_b Vectors
For Earthquakes and Explosions

	Station	m_b	M_s
Earthquakes			
	NP	5.90(.02)*	4.76(.04)
	WH	5.00(.04)	4.83(.06)
	PG	5.05(.03)	4.74(.04)
	RK	5.05(.02)	4.66(.04)
	HN	5.19(.02)	4.81(.04)
	KN	5.05(.04)	5.10(.05)
Presumed Explosions			
	NP	6.19(.07)	4.20(.10)
	WH	6.25(.40)	3.76(.60)
	PG	5.86(.06)	4.25(.08)
	RK	6.28(.04)	4.38(.07)
	HN	6.03(.05)	4.27(.07)
	KN	5.91(.07)	4.35(.10)

* One standard deviation of the mean.

The table indicates that the estimates for m_b have a smaller variance and that WH has a large variance associated with its estimated mean.

The test of the hypothesis that the overall earthquake and explosion vectors are equal yields a value for the Wilks generalized variance ratio $U_{2,1,112} = .335$ which can be compared with the 1% significance point $U_{2,1,110}(.01) = .911$, so that the hypothesis is rejected in accordance with expectations.

The comparison of earthquakes and explosions on a station by station basis generates $U_{2,6,112} = .256$ as compared with $U_{2,6,100}$ $(.01) = .772$. The values of the contrasts (i.e. the difference between the earthquake and explosion means on a station by station basis) are the most useful and we give simultaneous 95% confidence intervals for these contrasts in Table III. The 95% confidence interval for any contrast of the form $\underline{a}'B\underline{c}$ where $\underline{a}$ is q x 1 and $\underline{c}$ is p x 1, is given by (Morrison, 1970)

$$\underline{a}'\hat{B}\underline{c} \pm Q \cdot \left(\frac{q}{N-q} F_{\alpha:q,N-q}\right)^{1/2} \tag{25}$$

where

$$Q^2 = (\underline{c}' \hat{\Sigma} \underline{c}) \ (\underline{a}'(X'X)^{-1}\underline{a}) \tag{26}$$

TABLE III

Estimated Contrasts Between Earthquakes and Presumed Explosions

Difference by Station

	m_b	M_s	M_s-1.9m_b
NP	-.30 ± .08*	.56 ± .13	1.12 ± .13
WH	-.125 ± .19	1.07 ± .29	3.46 ± .30
PG	-.81 ± .07	.49 ± .13	2.03 ± .13
RK	1.22 ± .07	.28 ± .11	3.61 ± .11
HN	-.83 ± .08	.54 ± .12	2.12 ± .13
KN	-.86 ± .09	.75 ± .13	2.38 ± .14

* 95% Confidence Interval of the mean.

The values for the two terms in (22) can be read from the BMD-X-63 printout.

From Table III we see that all contrasts are significantly different from zero. An alternate proposed discriminant is the linear function represented symbolically as $M_s - 1.9\ m_b$. The values of this contrast are shown in Table III along with 95% confidence intervals. Again we see that all contrasts are significantly different from zero.

The fact that the mean (M_s, m_b) point for the earthquakes is different from that of the explosions is not, of course, solely indicative of discrimination capability since the means of either group can be shifted anywhere along its M_s, m_b trend by suitable selection of large or small events. However, examination of the M_s, m_b plots for the data shows that in fact the separation in the means is truly due almost exclusively to the separation of the trend lines. The fact that the earthquake trend line tends to have a slope between 1.6 and 1.9 leads to the discriminant proposed above.

Finally, we note that the main use of the estimates for the generalized mean vectors (22) and covariance matrix (23) will be in estimating the discriminant vectors $\underline{d}_j$ in (22) to be applied to each station (M_s, m_b) pair. The estimated discriminant vector,

$$\hat{\underline{d}}_j = (\hat{\underline{V}}_{1j} - \hat{\underline{V}}_{2j})' \hat{\Sigma}^{-1} \tag{28}$$

can be applied to a data sample of N_s observed station vectors to obtain the value (1)

$$\hat{d}(\underline{X}_1, \ldots \underline{X}_{N_s}) = \sum_{j=1}^{N_s} \hat{\underline{d}}_j' \underline{X}_j - \sum_j (CT)_j \tag{29}$$

to be compared with a threshold $C(\alpha)$ which produces a desired false alarm probability α. Of course, the rate α is only obtained in theory when the mean vectors and covariance matrices are known exactly. The explosion detection probability will be estimated by (8) with

$$\hat{D}_j^2 = 1/2(\hat{\underline{V}}_{1j} - \hat{\underline{V}}_{2j})'\Sigma^{-1}(\hat{\underline{V}}_{1j} - \hat{\underline{V}}_{2j}) \tag{30}$$

in place of D_j^2, (5). Since the explosion detection probability is an increasing function of $\sum_j D_j^2$, it is clear that D_j^2 measures the contribution to detection probability of station j. The overall detection probability can be increased either by an event which records at many stations or by an event which records at a relatively few stations with high D_j^2 values.

The above results can be illustrated using the data in Table I. A discriminant analysis program (DISNP) which accepts the generalized mean vectors and covariance

matrix along with each event and the identities of the stations recording each event was used. Table IV shows the resulting estimated discriminant vectors, correction terms and D^2 values for each station.

We note that station WH adds the most to the detection probability with NP adding the least. (WH may be good, however, mostly because of its small data sample, see Table I, an effect pointed out by Shumway and Blandford (1970)). In order to determine performance measured by plotting detection probability as a function of false alarm rate (Figure 1) we use equations (9) and (10) for D^2 = 1.7, 5.0, 10.0, and 15.0 in the program IDCURVE. All that is necessary to evaluate the performance of a particular sub-group of stations is to add up the D^2 values for the sub-group and refer to the identification curve in Figure 2. We see that the only politically unacceptable performance occurs when NP is the only station recording. Any two stations (exclusive of NP) which record give a D^2 of at least 10 yielding a detection probability of .90 for a false alarm probability .001. The worst combination of stations recording (PG, HN) in Table I still yield a D^2 of greater than 10 or a detection probability of greater than .90 in the allowable false alarm region. It is important to emphasize that the data set used consists of very large events and does not include any "anomalous" earthquakes such as those discussed by Der (1972).

The values of the discriminant vector applied to the 20 earthquakes and 9 explosions are shown in Table V.

If we arbitrarily assign earthquakes equal prior odds, the threshold is 0 and no misclassification result.

TABLE IV

Discriminant Analysis Results

STATION	DISCRIMINANT VECTOR	CORRECTION TERM	D^2
NP	(-5.02, 3.43)	-14.93	1.70
WH	(-16.37, 9.25)	-52.30	15.24
PG	(-9.81, 5.13)	-30.48	5.24
RK	(-13.03, 5.84)	-47.41	8.79
HN	(-10.18, 5.40)	-37.55	5.68
KN	(-11.25, 6.38)	-31.41	7.22

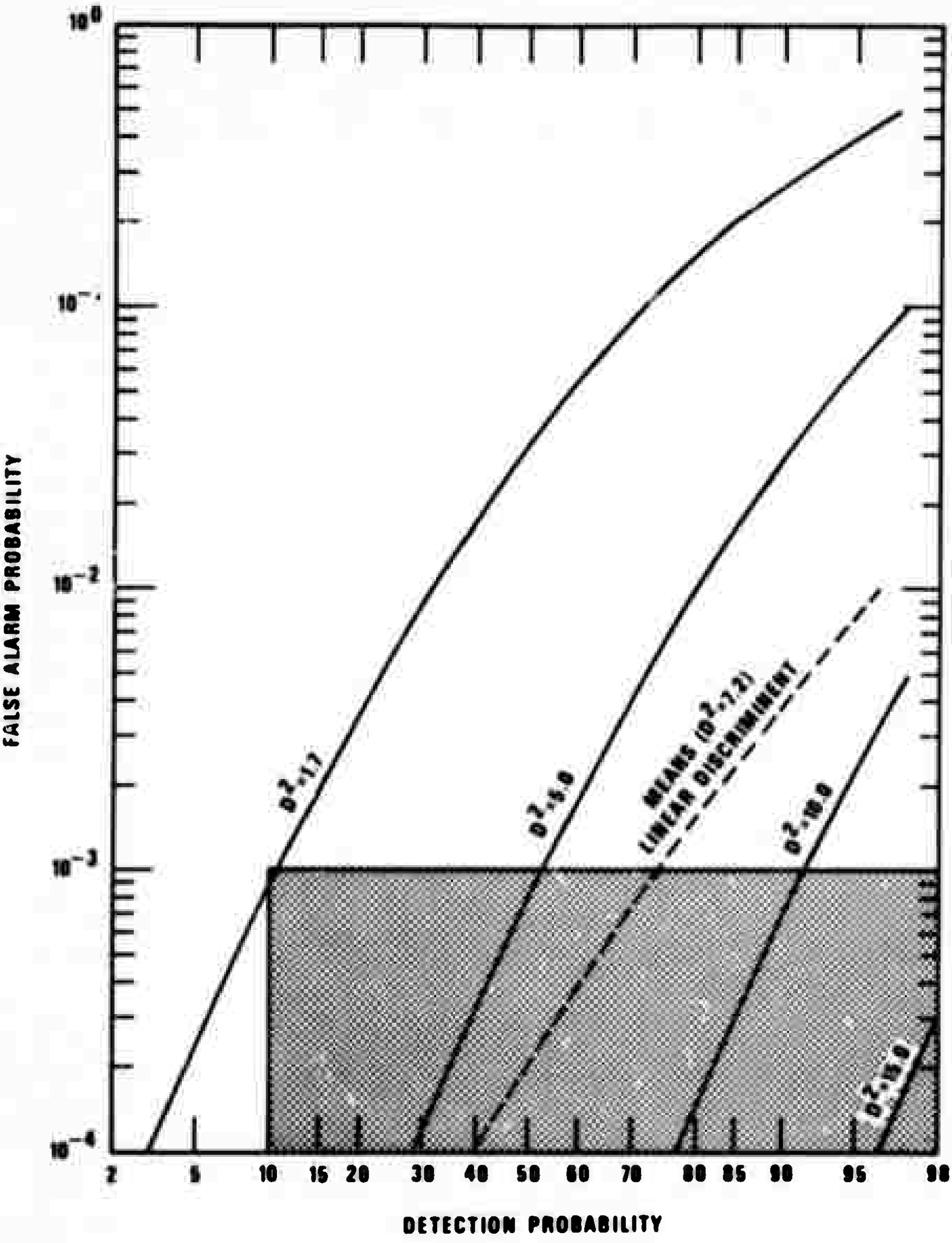

Figure 2. Identification curves for selected values of D^2 in multiple station discrimination.

TABLE V

Values of Discriminant Function (5.10)
For Data in Table I

Events	NP	WH	PG	RK	HN	KN	Discriminant Function
EQ1	X			X	X		2.50
EQ2	X			X	X		10.98
EQ3	X				X	X	7.42
EQ4	X			X	X		15.06
EQ5	X		X	X	X		34.15
EQ6	X		X	X	X		17.68
EQ7	X		X			X	19.87
EQ8	X	X	X	X		X	55.68
EQ9	X	X	X	X	X	X	56.76
EQ10	X			X	X		21.40
EQ11		X	X	X		X	13.70
EQ12	X	X	X	X	X	X	34.45
EQ13	X		X	X	X		10.23
EQ14	X	X	X	X	X		19.86
EQ15	X	X	X	X	X	X	51.69
EQ16	X	X	X	X	X	X	55.20
EQ17	X	X	X	X	X		27.45
EQ18		X	X	X			34.97
EQ19	X	X	X	X	X	X	63.68
EQ20		X	X	X	X	X	35.70
EX1				X	X	X	-21.03
EX2	X		X	X	X	X	-18.67
EX3	X		X	X		X	-11.62
EX4	X	X	X	X	X	X	-30.83
EX5				X	X		-12.35
EX6	X		X	X	X	X	-41.57
EX7			X	X	X		-23.17
EX8	X		X	X	X	X	-36.29
EX9	X		X	X	X		-34.47

It is interesting to compare the multiple station discriminant with one derived only from event means (the usual approach). Table VI shows the event means. In this case $N_s = 1$ and we obtain the discriminant vector (-15.20, 7.31) with correction term -52.45. The value of D^2 was 7.246, which puts the theoretical detection rate at .75 for a false alarm rate .001. Figure 2 shows the identification curve using the sample means and variances of the discriminant values rather than the theoretical means and variances based on D^2. Table VII shows the values for the discriminant function.

TABLE VI

Data on Measured Body Wave Magnitudes (m_b) (X_1) and Surface Wave Magnitudes (M_s) (X2) for 9 Presumed Explosions (pop. 2) and 20 Earthquakes (Pop. 1)

Explosions (E)		Earthquakes (Q)	
OBSERVATIONS		OBSERVATIONS	
POP2 m_b	M_s	POP1 m_b	M_s
6.04	4.33	5.60	4.25
5.97	4.39	5.18	3.93
5.85	4.35	6.31	6.30
5.79	4.14	5.36	4.49
5.87	3.90	5.96	6.39
6.51	4.49	5.26	4.42
5.74	4.22	5.17	5.10
5.98	4.08	4.75	4.50
6.07	4.30	5.35	5.49
		5.01	4.48
		5.27	4.41
		5.27	4.69
		4.98	3.66
		5.11	3.99
		5.06	4.58
		5.09	4.90
		5.15	4.82
		4.56	4.08
		5.00	4.94
		5.43	5.48

TABLE VII

Values of Discriminant Function for Data in Table VI

EQ1	4.61	EX1	-7.71
EQ2	2.44	EX2	-6.21
EQ3	2.58	EX3	-4.68
EQ4	3.79	EX4	-5.30
EQ5	8.56	EX5	-8.27
EQ6	4.80	EX6	-13.69
EQ7	11.14	EX7	-3.96
EQ8	13.14	EX8	-8.36
EQ9	11.25	EX9	-8.39
EQ10	9.04		
EQ11	4.58		
EQ12	6.62		
EQ13	3.50		
EQ14	3.94		
EQ15	9.01		
EQ16	10.89		
EQ17	9.40		
EQ18	12.96		
EQ19	12.55		
EQ20	9.96		

6. OTHER POSSIBLE APPROACHES, SUMMARY

Two other approaches were tried for the multiple station case. The first was to use the (M_s, m_b) pairs from the six stations as a 12-element vector with missing observations filled in as sample means. This tended to produce a 12 x 12 covariance matrix which was nearly singular. Since some columns became nearly linear combinations of the others this tended to produce a 12 x 12 covariance matrix which was nearly singular. Thus the performance of the discriminant on future data as in Shumway and Blandford (1970) could not be evaluated, and doubt was cast on the entire technique. A second approach used a non-linear function composed of the ratio of generalized variances under the assumption that the suite of station observations originated from either a bomb or an earthquake. While this procedure was appealing as the generalized likelihood solution, the test statistic was not normally distributed. Detection curves based on a normal approximation were inferior to those for the multiple station linear discriminant.

The preceding calculations indicate that using a multiple station discriminant function when the appropriate data is available will lead to an average overall improvement in the identification curves. Consideration of the results in Shumway and Blandford, coupled with the realization that our new technique is still basically two-dimensional instead of 12-dimensional, suggests that the improvement is real and not due to higher dimensionality of the new technique working on a small data base. However this possibility is not definitely excluded in this study

which is aimed more at outlining the technique and giving an illustrative example.

This improvement may be largely due to the proper accounting made of "station effects" which would be the same for earthquakes and explosions and which could also be accounted for by the traditional method of "station corrections" applied before calculation of the average M_s, m_b values. The present method, while preserving the capability of discrimination when reports from some stations are missing, can in addition take advantage of different M_s or m_b patterns between earthquakes and explosions, i.e. the possibility of different radiation patterns.

Several possible additional sources of improvement need to be investigated. These improvements could be possible using more detailed assumptions about the covariance matrices. For example, it is generally thought that the scatter for the earthquake distribution exceeds that of the explosion distribution although the performance of the classical discriminant is not degraded significantly by this violation (Shumway and Blandford, 1970). In any case, a class of linear functions of the form

$$d_j' = (V_{1j} - V_{2j})'(\Sigma_1 - \theta\Sigma_2)^{-1}$$

with Σ_1, Σ_2 the explosion and earthquake covariance matrices and θ a parameter chosen to optimize the detection probability might make a significant improvement. A second possibility is to take advantage of

inhomogeneities in the covariance matrices for different stations by replacing Σ in equation (2) by Σ_j.

A problem of practical interest arises when a few, but not all, discriminant measures are missing from a station. In this case, one still knows the marginal distribution of the variables present to be normal and the analysis proceeds as before. Let Σ be partitioned into

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

where all values corresponding to the second set of p_2 variables are missing. Then the marginal density of the first set of p_1 present variables is normal with covariance matrix Σ_{11}. Hence, in the discriminant function (2) just replace Σ by Σ_{11} and $(V_{1j} - V_{2j})$ by $(V^*_{1j} - V^*_{2j})$ where the starred mean difference is a p_1 by 1 vector composed of the first p_1 components of $(V_{1j} - V_{2j})$. In these cases, one still gets the use of a station value as a possible discriminator even though it is a vector of lower dimension than one might prefer under ideal conditions. This requires that one retain a different covariance matrix for each station but preserves the optimality properties that would be destroyed by guessing values for the missing variables.

7. ACKNOWLEDGEMENT

The authors are indebted to H. Husted for computer assistence and to D. Nelson for calculating the M_s, m_b values.

8. REFERENCES

Booker, A. and Mitronovas, W., 1964, An application of statistical discrimination to classify seismic events: Bull. Seism. Soc. Am., v. 54, p. 961-977.

Capon, J., Greenfield, R.J., and Lacoss, R.T., 1967, Surface vs body wave magnitude results: Seismic Discrimination Semiannual Technical Summary, 30 June 1967, Lincoln Laboratory, Lexington, Massachusetts.

Der, Z., 1972, M_s - m_b characteristics of earthquakes in the Eastern Himalayan regions: Seismic Data Laboratory Report No. 296, Teledyne Geotech, Alexandria, Virginia.

Dixon, W.J., 1970, BMD biomedical computer programs, X-series supplement: University of California Press, Berkeley, California.

Ericsson, Ulf A., 1970, Event identification for test ban control: Bull. Seism. Soc. Am., v. 60, No. 5, p. 1521-1546.

Morrison, D.L., 1970, Multivariate techniques, McGraw Hill.

Shumway, R.H. and Blandford, R.R., 1970, A simulation of seismic discriminant analysis: Seismic Data Laboratory Report No. 261, Teledyne Geotech, Alexandria, Virginia.